IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

First Named Inventor :	Yunis Mohammed	Appeal No. ---
Appln. No. :	09/751,871	
Filed :	December 29, 2000	Group Art Unit: 2641
For :	COMPRESSED LEXICON AND METHOD OF APPARATUS FOR CREATING AND ACCESSING THE LEXICON	Examiner: Michael N. Opsasnick
Docket No.:	M61.12-0334	

BRIEF FOR APPELLANT

Electronically Filed
June 23, 2008

Mail Stop Appeal Brief-Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450
Sir:

This is an appeal from an Office Action dated September 17, 2007 in which claims 1-19 and 22 were finally rejected.

REAL PARTY IN INTEREST

Microsoft Corporation, a corporation organized under the laws of the state of Washington, and having offices at One Microsoft Way, Redmond, Washington 98052, has acquired the entire right, title and interest in and to the invention, the application, and any and all patents to be obtained therefor, as set forth in the Assignment filed with the patent application and recorded on Reel 011429, frame 0162.

RELATED APPEALS AND INTERFERENCES

There are no known related appeals or interferences which will directly affect or be directly affected by or have a bearing on the Board's decision in this appeal.

STATUS OF THE CLAIMS

I. Total number of claims in the application.

Claims in the application are: 1-31

II. Status of all the claims.

A.	Claims cancelled:	20-21 and 23-31
B.	Claims withdrawn but not cancelled:	0
C.	Claims pending:	1-19 and 22
D.	Claims allowed:	0
E.	Claims rejected:	1-19 and 22
F.	Claims Objected to:	0

III. Claims on appeal

The claims on appeal are: 1-19 and 22

STATUS OF AMENDMENTS

There are no outstanding amendments.

SUMMARY OF CLAIMED SUBJECT MATTER

Of the claims currently on Appeal, there are three independent claims: claim 1, claim 12 and claim 19. Claim 1 is a method of building a compressed speech lexicon for use in a speech application 202. In building the compressed speech lexicon, a system first receives a word list configured for use in the speech application 202, the word list including a plurality of words 235, with each word in the word list having associated word-dependent data selected from the group consisting of pronunciation data 238 and part-of-speech data 240. See specification, page 14, lines 7-24. See also FIG. 3 and FIG. 4. Next, one of the words 235 is selected, and an index entry is generated. The index entry identifies a location in a compressed speech lexicon memory 226 for holding the selected word 235. See steps 250-262 in FIG. 4, and specification, page 15, lines 3-27, page 16, line 19-page 17, line 21. The system then encodes the selected word 235 and its associated word-dependent data to obtain encoded words and associated encoded word-dependent data. See block 274 in FIG. 4, and specification page 17, line 22-page 18, line 8. Finally, the encoded word and its associated word-dependent data are written to the identified location in the speech lexicon memory 228. See block 276 in FIG. 4 and specification page 18, lines 9-16.

Another set of claims is drawn to a method of accessing word information related to a word stored in a compressed speech lexicon. This is set out in independent claim 12. The

method includes first receiving the word and accessing an index (such as a hash table 232) to obtain a word location in a compressed speech lexicon (in memory 226) that contains information associated with the received word, including word-dependent data selected from the group consisting on pronunciation data 238 and part-of-speech data 240. See FIG. 5 and blocks 310-316 in FIG. 6, and specification, page 21, line 7-page 22, line 14. The encoded word information (290-296') is read from the word location. See FIGS. 5 and 6, specification page 22, line 15-page 23, line 6. The word information is then decoded for use in a speech application. See FIGS. 5 and 6, specification page 23, line 26-page 24, line 11.

A third set of claims is drawn to a compressed speech lexicon builder 220 for building a compressed speech lexicon for use in a speech application 202 based on a word list containing a plurality of domains. This is set out in claim 19. The domains include words 235 and word-dependent data associated with each of the words. The compressed speech lexicon builder comprises a plurality of domain encoders 268, 270 and 272. One of the domain encoders is associated with each domain in the word list. See FIG. 3, page 13, line 23-line 29. The domain encoders 268-272 are configured to compress the words and the associated word-dependent data that is selected from the group consisting of pronunciation data 236 and part-of-speech data 240. See FIG. 3, specification page 14, line 17-line 24. The compressed lexicon builder 220 also includes a hashing component 224 that generates a hash value 256 for each word 235 in the word list. A hash table generator 252 determines the next available location in a speech lexicon memory 226 and writes, at an address in the hash table 232 identified by the hash value 256, the next available location in the speech lexicon memory 228. See specification, page 16, line 19-page 17, line 12. A speech lexicon memory generator 246 stores in the speech lexicon memory 228 compressed words and compressed word-dependent data, for use by the speech application 202. See FIG. 3, specification page 12, line 22-page 13, line 16. Each compressed word and its associated compressed word-dependent data are stored in the next available location in the speech lexicon memory 228 written in the hash table 232 at the hash table address associated with the compressed word. See FIG. 5, specification page 18, lines 9-16.

The claims do not stand or fall together. Instead, they are grouped as follows:

Claims 1-11;

Claims 12-18; and

Claims 19 and 22.

GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

The grounds of rejection to be reviewed on appeal are whether claims 1-11, 19 and 22 are obvious under 35 U.S.C. §103(a) in view of Burrows US Patent No. 6,021,409, in view of Sarukkai et al. US Patent No. 5,819,220 and further in view of Poirer et al. US Patent No. 6,321,372; and

Whether claims 12-18 are obvious under 35 U.S.C. §103(a) as being unpatentable over Burrows US Patent No. 6,021,409 in view of Pringle et al. US Patent No. 6,470,306, and further in view of Poirer et al. US Patent No. 6,321,372.

ARGUMENT

I. Claims 1-11 are patentable over Burrows in view of Sarukkai et al. and further in view of Poirer et al.

On page 2 of the final Office Action in the above-identified matter, the Examiner rejected claims 1-11 under 35 U.S.C. §103(a) as being unpatentable over Burrows US Patent No. 6,021,409, in view of Sarukkai et al. US Patent No. 5,819,220 and further in view of Poirer et al. US Patent No. 6,321,372. Appellant asserts that this rejection, made in the final Office Action, should be reversed. Claim 1 is directed to a method of building a compressed lexicon for use in a speech application. In order to have an effective speech application (such as a text-to-speech or speech synthesis system; or such as a speech recognition system) a relatively large vocabulary lexicon is required. The lexicon can contain a word list and word-dependent data, such as pronunciation information and part-of-speech information. The large vocabulary lexicon may be desirable in order to ensure that the coverage of the speech application is adequate (in order to ensure that it can address substantially all words that it will encounter). However, such a large lexicon can be difficult to maintain, and cumbersome to use in a speech application.

Similarly, it may be desirable that the speech synthesis or speech recognition results

be provided by the application very quickly. Therefore, while compression of data may address some disadvantages associated with a very large lexicon, data compression brings its own disadvantages. For instance, many compression algorithms make it cumbersome to recover the compressed data. This can result in an undesirable amount of processing time to return a result, especially with respect to the desired time limitations imposed on speech recognition and speech synthesis tasks.

Therefore, claim 1 is directed to a method of building a speech lexicon for use in a speech application in a way in which both the encoded word, and its encoded word-dependent data, can be quickly accessed so that the performance of the speech application is not compromised.

In rejecting independent claim 1 the Examiner asserted that Burrows discloses all of the elements of the claim except that Burrows "teaches the use of the word techniques in an internet environment" rather than "using the word techniques in a speech related application". However, the Examiner found Sarukkai et al. to disclose "using word list techniques in web based speech applications" and concluded that it would have been obvious to adapt the teachings of Burrows into speech related web applications "because it would advantageously tailor the speech enabled sites to specific vocabularies . . . ."

The Examiner also stated that the combination of Burrows and Sarukkai et al. fail to disclose the use of the words in a speech lexicon memory. However, the Examiner asserted that Poirer et al. discloses "the providing of internet information in the form of providing linguistic services that include speech lexicon" and concluded that it would have been obvious to "modify the teachings of the combination of Burrows . . . in view of Sarukkai et al. . . . with the use of speech lexicons because it would advantageously be used to provide linguistic services . . . ."

Appellant respectfully believes that there is no motivation to combine the cited references because they bear no relation to each other or the present invention, there is no disclosure in Sarukkai et al. or Poirer et al. relating to any method or technique for transforming the web page indexing method of Burrows to form the invention of claim 1, and the cited combination fails to disclose all of the elements of the claims.

Burrows discloses a method of parsing and indexing a web page. The subject matter of Burrows has no relation to the present invention. In particular, there are fundamental differences between a speech lexicon and other structures that have some surface similarities to a lexicon. For example, a speech lexicon in the present context contains information related to the pronunciation and/or recognition of a spoken word. This word-dependent information is clearly lacking from the Burrows reference. Further, the Burrows reference makes no mention of word-dependent data for use in speech recognition. Burrows treats items that are not words, but information, such as metawords, as separate words that are indexed along with words parsed from a web page. Thus, the Burrows reference does not disclose anything related to a speech lexicon, nor does it disclose word-dependent data as provided in the claims.

While the Sarukkai et al. reference is directed to a computer system for user-provided speech actuation and access to stored information, it simply fails to teach or suggest, or remedy in any way, the deficiencies of the Burrows reference. Sarukkai et al. describe the basics of speech recognition and a method for dealing with out-of-context words. However, while the Sarukkai et al. reference may relate to speech recognition and web applications as suggested by the Examiner, the reference has nothing whatsoever to do with generating a compressed speech lexicon for use in a speech application. Further, the cited section of Sarukkai et al. (col. 3, lines 39-45) reads as follows:

> It is infeasible to build RGDAGs characterizing very large vocabulary spontaneous speech. Accordingly, a need remains for a better form of speech interface to a computer, in particular, one that is adapted to the broad vocabulary encountered on the World Wide Web, but able to respond to short speakable sequences of words not necessarily found in the large textual base of web pages and other Web-accessed documents.

Nothing in the cited section teaches how to modify the web page indexing method of Burrows to one tailored to speech related applications or provides any motivation for such a modification, as suggested by the Examiner. Therefore, the Appellant respectfully submits that the Sarukkai reference is also inapplicable to the present set of claims, and in any case, does not teach or suggest the limitations set out in the present claims, including the claimed word-dependent data.

The cited section of Poirer et al. (col. 8, lines 52-65) reads as follows:

> A "linguistic service" is a service that relates to one or more natural languages. Therefore, the broad scope of linguistic services encompasses any language-related operation that a user might request. Examples include tokenization, morphological analysis, part-of-speech tagging or disambiguation, low-level pattern extraction, stemming and lemmatizing, language identification, optical character recognition (OCR), speech recognition, dictionary and lexicon lookup, translation assistance, text extraction, summarization, annotation and glossing, information retrieval, shallow parsing, comprehension assistance, language-related knowledge management, indexation, idiom recognition noun phrase extraction, verb phrase extraction, and various combinations of these services.

However, there is no teaching in the cited section as to how one would modify the web page word indexing method of Burrows as modified by Sarukkai et al. such that the indexed words are used to build a compressed speech lexicon. The mere mention of "lexicon" in Poirer et al. is clearly insufficient to support a *prima facie* case of obviousness against the claims.

The cited, unrelated references fail to disclose the significant modifications to the web page word index of Burrows, as suggested by the Examiner, that would be required to form the present invention as described in independent claim 1. This is primarily due to the fact that none of the references is related to building a compressed speech lexicon for use in a speech application. Thus, not only do the references fail to disclose the claimed invention, but there is no motivation to combine the references in the first place, outside of Appellant's disclosure.

In the last Amendment, Appellant amended claim 1 to further clarify that the invention deals with a compressed speech lexicon for use in a speech application, to which the primary references cited by the Examiner do not relate. In particular, independent claim 1 was amended to recite "receiving a word list configured for use in the speech application, the word list including a plurality of words, each word in the word list having associated word-dependent data selected from the group consisting of a pronunciation and a part-of-speech . . ."

The cited word list of Burrows (col. 6 line 60-67) is unrelated to a word list "configured for use in the speech application" as provided in claim 1. Moreover, the cited word list of

Burrows (i.e., word indexed web pages) does not include "associated word-dependent data selected from the group consisting of a pronunciation and a part-of-speech" for each word in the word list. Rather, the word list of Burrows merely contains a list of the words found in pages 200 that are returned to the browser 20 in response to the request 21. The cited word list of Burrows does not contain a pronunciation or a part of speech for each of the words because the cited word list is unrelated to a word list that is configured for use in a speech application.

Additionally, there is no disclosure in Sarukkai et al. or Poirer et al. of the claimed word list and word-dependent data. Further, neither Sarukkai et al. nor Poirer et al. suggest or teach how one would modify the cited word list of Burrows to be configured for use in the speech application or how one would modify the word list of Burrows to include the claimed word-dependent data.

Accordingly, claim 1 is non-obvious is view of the cited references because the references fail to disclose all of the claimed elements. Additionally, claims 2-11 are allowable in view of the cited references at least due to their dependence from allowable base claim 1. Therefore, Appellant requests that the rejections of claims 1-11 be reversed.

II. Claims 12-18 are patentable over the combination of Burrows in view of Pringle et al. and further in view of Poirer et al.

In section 4 of the Office Action, the Examiner rejected claims 12-18 under 35 U.S.C. §103(a) as being unpatentable over Burrows in view of Pringle et al. (US Patent No. 6,470,306) and further in view of Poirer et al. Appellant respectfully submits that the rejection should be reversed.

Independent claim 12 is directed to a method of accessing word information related to a word stored in a compressed speech lexicon. In rejecting claim 12, the Examiner asserted that Col. 5, lines 15-35, and Col. 6, lines 17-42 of Burrows disclose all of the elements of the claim except for using the apparatus of Burrows for speech lexicon applications or the use of words in a speech lexicon memory. However, the Examiner found these deficiencies of Burrows to be

overcome by the teachings of Pringle et al. and Poirer et al. Appellant respectfully disagrees with the Examiner's assessment of the cited references.

Both the Burrows and the Poirer et al. references are unrelated to a compressed speech lexicon and accessing word information related to a word stored in a compressed speech lexicon. The Pringle reference also has nothing to do with a compressed speech lexicon. The Pringle reference is directed to a method and apparatus for translating a document from one language to another language. This is commonly referred to as machine translation and involves natural language processing, but not necessarily speech recognition. The Pringle reference simply does not disclose anything related to a speech application or a compressed speech lexicon. In fact, the process disclosed in Pringle would not be an acceptable process for use on a speech lexicon. Therefore, the Pringle reference simply cannot teach or suggest the present claims, either alone or in combination with any of the other references cited by the Examiner.

More importantly, Pringle et al. fail to teach or suggest the modification of a word index of web pages could be modified or otherwise incorporated into a speech translation system at Col. 2, lines 40-60, as suggested by the Examiner. The cited section provides:

> The automated natural language translation system according to the invention has many advantages over known machine-based translators. After the system of the invention automatically selects the best possible translation of the input textual information and provides the user with an output (preferably a Japanese language or Spanish language translation of English-language input text), the user can then interface with the system to edit the displayed translation or to obtain alternative translations in an automated fashion. An operator of the automated natural language translation system of the invention can be more productive because the system allows the operator to retain just the portion of the translation that he or she deems acceptable while causing the remaining portion to be retranslated automatically. Since this selective retranslation operation is precisely directed at portions that require retranslation, operators are saved the time and tedium of considering potentially large numbers of incorrect, but highly ranked translations. Furthermore, because the system allows for arbitrary granularity in translation adjustments, more of the final structure of the translation will usually have been generated by the system. The system thus reduces the potential for human (operator) error and saves time in edits that may involve structural, accord, and tense changes. The system efficiently gives operators the fill benefit of its extensive and reliable knowledge of grammar and spelling.

Accordingly, the teaching or suggestion cited by the Examiner as being taught by Pringle et al. is erroneous.

Moreover, none of the cited references discloses "accessing an index to obtain a word location in the compressed speech lexicon that contains information associated with the received word including word-dependent data selected from the group consisting of a pronunciation and a part-of-speech" as provided in claim 12. As mentioned above, Burrows, being unrelated to speech recognition applications, fails to include any word-dependent data used by speech recognition applications including either a pronunciation or a part-of-speech. Further, neither Pringle et al. nor Poirer et al. disclose the claimed speech lexicon or a method step of accessing an index to obtain a word location in the speech lexicon, as provided in claim 12.

Accordingly, claim 12 is non-obvious in view of the cited references because there is no motivation to combine the references outside of Appellant's disclosure and they fail to disclose

all of the claimed elements. Additionally, claims 13-18 are allowable in view of the cited references at least due to their dependence on allowable claim 12. Therefore, Appellant requests that the rejections of claims 12-18 be reversed.

III. Claims 19 and 22 are patentable over Burrows in view of Sarukkai et al. and further in view of Poirer et al.

In section 3 of the Office Action, the Examiner rejected claims 19 and 22 under 35 U.S.C. §103(a) as being unpatentable over Burrows, in view of Sarukkai et al., and further in view of Poirer et al. Appellant respectfully traverses the rejection and submits that it should be reversed.

Claim 19 is directed to a compressed speech lexicon builder for building a compressed speech lexicon for use in a speech application. The Examiner rejected claim 19 for the same reasons as claim 1. For the same reasons as claim 1, the references fail to teach or suggest claim 19. Also, for the same reasons discussed above with respect to claim 1, Appellant respectfully submits that there is no motivation to combine the cited references because they bare no relation to each other or the present invention.

Further, claim 19 recites "a plurality of domain encoders, one domain encoder being associated with each domain in the word list, the domain encoders being configured to compress the words and the associated word-dependent data selected from the group consisting of a pronunciation and a part-of-speech, to obtain compressed words and compressed word-dependent data."

In rejecting claim 19, the Examiner asserted that Burrows teaches "'a compressed lexicon….builder' as word list with domain such as attributes (Col. 9 lines 21-29)". The cited section of Burrows provides:

attributes.

For example, the page **200** of FIG. **4** can have associated page attributes **250**. Page attributes **250** can include □ADDRESS□ **251**, □DESCRIPTION□ **252**, □SIZE□ **253**, □DATE□ **254**, □FINGERPRINT□ **255**, □TYPE□ **256**, and □END_PAGE□ **257**, for example. The symbol "□," represents one or more characters which cannot be confused with the characters normally found in words, for example "space," "underscore," and "space" (sp_sp).

Clearly, the cited section of Burrows fails to disclose the compressed speech lexicon builder of claim 19 "for building a compressed speech lexicon for use in a speech application based on a word list containing a plurality of domains, the domains including words and word-depend data associated with each of the words". Rather, the cited compressed lexicon builder of Burrows is unrelated to a speech application and the cited attributes do not represent word-dependent data corresponding to that used for speech applications.

Also in rejecting claim 19, the Examiner found Burrows to teach "'a plurality of domain encoders….data' as compressing the word entries based on delta values (Col. 11, line 40-col. 12 line 26)". However, the cited section of Burrows merely discloses a "prefix compressing technique which can be used to map from word 710 to compressed word 720" and "a delta value compressing technique which can be applied to the locations 800 of FIG. 6 . . . and takes advantage of the fact that frequently occurring words such as 'the', 'of', 'in', etc. are close to each other." [Col. 11, lines 41-42 and Col. 11, lines 59-62] Appellant cannot discern the elements of Burrows that correspond to the claimed "plurality of domain encoders" from the information provided by the Examiner. Even if one believes that the cited section of Burrows discloses domain encoders, neither Burrows nor any of the other cited references disclose that such encoders "compress the words and the associated word-dependent data selected from the group consisting of a pronunciation and a part of speech, to obtain compressed words [of the plurality of domains] and compressed word-dependent data" as provided in claim 19.

Accordingly, Appellant believes that claim 19 is non-obvious in view of the cited references because there is no motivation or suggestion for combining the references outside of Appellant's disclosure and the references fail to disclose all of the claimed elements.

Additionally, claim 22 is non-obvious in view of the cited references at least due to its dependence on allowable claim 19. Therefore, Appellant requests that the rejections of claims 19 and 22 be reversed.

III. Conclusion

In view of the arguments provided above, Appellants assert that claims 1-19, and 22, are patentable over the references cited by the Examiner. Therefore, Appellant requests that the rejection of claims 1-19 and 22 be reversed and that claims 1-19 and 22 be allowed.

Respectfully submitted,

WESTMAN, CHAMPLIN & KELLY, P.A.

By: /Joseph R. Kelly/

Joseph R. Kelly, Reg. No. 34,847
900 Second Avenue South, Suite 1400
Minneapolis, Minnesota 55402-3319
Phone:(612) 334-3222 Fax:(612) 334-3312

JRK:slg

Appendix A

1. A method of building a compressed speech lexicon for use in a speech application, comprising:

receiving a word list and word-dependent data, configured for use in the speech application, associated with each word in the word list;

selecting a word from the word list;

generating an index entry identifying a location in a compressed speech lexicon memory for holding the selected word;

encoding the selected word and its associated word-dependent data to obtain encoded words and associated encoded word-dependent data; and

writing the encoded word and its associated word-dependent data at the identified location in the speech lexicon memory.

2. The method of claim 1 and further comprising:

repeating the steps of selecting, generating, encoding and writing for each word in the word list and the associated word-dependent data.

3. The method of claim 2 and further comprising:

writing codebooks corresponding to the encoded words and the encoded word-dependent data in the speech lexicon memory.

4. The method of claim 1 wherein receiving the word list comprises:

counting the words in the word list;

allocating a hash table memory based on a number of words in the word list; and

allocating a speech lexicon memory based on the number of words in the word list.

5. The method of claim 1 wherein generating an index entry comprises:

determining a next available location in the speech lexicon memory.

6. The method of claim 5 wherein generating an index entry comprises:

calculating a hash value for the selected word;

indexing into the hash table to an index location based on the hash value; and

writing location data identifying the next available location in the speech lexicon memory into the index location in the hash table.

7. The method of claim 6 wherein writing location data comprises:

writing an offset into the speech lexicon memory that corresponds to the next available location in the speech lexicon memory.

8. The method of claim 1 wherein encoding comprises:

providing a word encoder to encode the words in the word list and encoding the words with the word encoder; and

providing word-dependent data encoders for each type of word-dependent data in the word list and encoding the word-dependent data with the word-dependent data encoders.

9. The method of claim 8 wherein encoding further comprises:

Hufmann encoding the selected word and its associated word-dependent data.

10. The method of claim 1 wherein writing the encoded word and word-dependent data comprises:

writing a data structure comprising:

a word portion containing the encoded word;

a word-dependent data portion containing the encoded word-dependent data; and

wherein each word-dependent data portion has an associated last indicator portion

and word-dependent data indicator portion, the last indicator portion containing an indication of a last portion of word-dependent data associated with the selected word, and the word-dependent data indicator portion containing an indication of the type of word-dependent data stored in the associated word dependent data portion.

11. The method of claim 10 wherein writing a data structure comprises writing the word portion and the word-dependent data portions as variable length portions followed by a separator.

12. A method of accessing word information related to a word stored in a compressed speech lexicon, comprising:

receiving the word;

accessing an index to obtain a word location in the compressed speech lexicon that contains information associated with the received word;

reading encoded word information from the word location; and

decoding the word information for use in a speech application.

13. The method of claim 12 and further comprising:

prior to reading the encoded word information, reading an encoded word from the word location;

decoding the encoded word; and

verifying that the decoded word is the same as the received word.

14. The method of claim 12 wherein reading the encoded word information comprises:

reading a plurality of fields from the word location containing variable length word information.

15. The method of claim 14 wherein reading a plurality of fields comprises:

prior to reading each field, reading data type header information indicating a type of word information in an associated field.

16. The method of claim 15 wherein reading a plurality of fields comprises:
reading a last field indicator indicating whether an associated one of the plurality of fields is a last field associated with the received word.

17. The method of claim 12 wherein decoding the word information comprises:
initializing decoders associated with the word and its associated information.

18. The method of claim 12 wherein accessing an index comprises:
calculating a hash value based on the received word;
finding an index location in the index based on the hash value; and
reading from the index location a pointer value pointing to the word location in the compressed lexicon.

19. A compressed speech lexicon builder for building a compressed speech lexicon for use in a speech application based on a word list containing a plurality of domains, the domains including words and word-dependent data associated with the words, the compressed speech lexicon builder comprising:
a plurality of domain encoders, one domain encoder being associated with each domain in the word list, the domain encoders being configured to compress the words and word-dependent data to obtain compressed words and compressed word-dependent data;
a hashing component configured to generate a hash value for each word in the word list;
a hash table generator, coupled to the hashing component, configured to determine a next available location in a speech lexicon memory and write, at an address in a hash table identified by the hash value, the next available location in the speech lexicon

memory; and

a speech lexicon memory generator, coupled to the domain encoders and the hash table generator, configured to store in the speech lexicon memory, for use by the speech application, the compressed words and compressed word-dependent data, each compressed word and its associated compressed word-dependent data being stored at the next available location in the speech lexicon memory written in the hash table at the hash table address associated with the compressed word.

20. Canceled.

21. Canceled.

22. The compressed speech lexicon builder of claim 19 and further comprising:

a codebook generator generating a codebook associated with each domain encoder.

23. Canceled.

24. Canceled.

25. Canceled.

26. Canceled.

27. Canceled.

28. Canceled.

29. Canceled.

30. Canceled.

31. Canceled.

EVIDENCE APPENDIX

None.

RELATED PROCEEDINGS APPENDIX

None.